Online Appendix for "Forecasting the Equity Risk Premium: The Role of Technical Indicators"

Christopher J. Neely
Federal Reserve Bank of St. Louis
neely@stls.frb.org

David E. Rapach
Saint Louis University
rapachde@slu.edu

Jun Tu
Singapore Management University
tujun@smu.edu.sg

Guofu Zhou
Washington University in St. Louis and CEMA
zhou@wustl.edu

August 20, 2013

This Online Appendix describes the wild bootstrap procedure used to compute empirical p-values in Section 2 of the paper. It also reports the complete structural stability test results (Table A1) discussed in footnote 9 of the paper, as well as the predictive regression results for ΔSENT (Table A2) and conditional asset pricing model estimation results (Table A3) discussed in Section 2.4 of the paper.

We begin with the wild bootstrap procedure. Let

$$\hat{\varepsilon}_{t+1} = r_{t+1} - (\hat{\alpha} + \sum_{i=1}^{N} \hat{\beta}_{i,x} x_{i,t} + \sum_{i=1}^{N} \hat{\beta}_{i,S} S_{i,t}), \tag{A1}$$

where $\hat{\alpha}$, $\hat{\beta}_{i,x}$ $(i = 1, \ldots, N)$, and $\hat{\beta}_{i,S}$ $(i = 1, \ldots, N)$ are OLS parameter estimates for the general multiple predictive regression model that includes a constant and all of the N macroeconomic variables and N technical indicators as regressors. Following convention, we assume that each macroeconomic variable follows an AR(1) process:

$$x_{i,t+1} = \rho_{i,0} + \rho_{i,1} x_{i,t} + v_{i,t+1} \text{ for } i = 1, \ldots, N. \tag{A2}$$

Define

$$\hat{v}^c_{i,t+1} = x_{i,t+1} - (\hat{\rho}^c_{i,0} + \hat{\rho}^c_{i,1} x_{i,t}) \text{ for } i = 1, \ldots, N, \tag{A3}$$

where $\hat{\rho}^c_{i,0}$ and $\hat{\rho}^c_{i,1}$ are reduced-bias estimates of the AR(1) parameters in (A2). The reduced-bias estimates of the AR parameters are computed by iterating on the analytical second-order bias expression for the OLS estimates. Based on these AR parameter estimates and fitted residuals, we build up a pseudo sample of observations for the equity risk premium and N macroeconomic variables under the null hypothesis of no return predictability:

$$r^*_{t+1} = \bar{r} + \hat{\varepsilon}_{t+1} w_{t+1} \text{ for } t = 0, \ldots, T-1, \tag{A4}$$

$$x^*_{i,t+1} = \hat{\rho}^c_{i,0} + \hat{\rho}^c_{i,1} x^*_{i,t} + \hat{v}^c_{i,t+1} w_{t+1} \text{ for } i = 1, \ldots, N \text{ and } t = 0, \ldots, T-1, \tag{A5}$$

where $\bar{r}$ is the sample mean of the equity risk premium, w_{t+1} is a draw from the standard normal distribution, and $x^*_{i,0} = x_{i,0}$ $(i = 1, \ldots, N)$. Observe that we multiply $\hat{\varepsilon}_{t+1}$ in (A4) and each $\hat{v}^c_{i,t+1}$ in (A5) by the same scalar, w_{t+1}, when generating the month-$(t+1)$ pseudo residuals, thereby making it a wild bootstrap. In addition to preserving the contemporaneous correlations in the data, the wild bootstrap accounts for general forms of conditional heteroskedasticity. Employing reduced-bias parameter estimates in (A5) helps to ensure that we adequately capture the persistence in the macroeconomic variables. To generate a pseudo sample of observations for the N technical indicators, we assume that each indicator follows a first-order, two-state, Markov-switching process with the following transition matrix:

$$P_i = \begin{pmatrix} p_i^{0,0} & p_i^{1,0} \\ p_i^{0,1} & p_i^{1,1} \end{pmatrix} \text{ for } i = 1, \ldots, N, \tag{A6}$$

where

$$p_i^{j,k} = \Pr(S_{i,t} = k | S_{i,t-1} = j) \text{ for } j, k = 0, 1, \tag{A7}$$

and $p_i^{0,0} + p_i^{0,1} = p_i^{1,0} + p_i^{1,1} = 1$. Based on estimates of the transition probabilities in (A6) and $S_{i,0}$ $(i = 1, \ldots, N)$, we can build up a pseudo sample of observations for the N technical indicators via simulations.

Using the pseudo sample of observations for the equity risk premium $[\{r^*_{t+1}\}_{t=0}^{T-1}]$, N macroeconomic variables $[\{x^*_{i,t}\}_{t=0}^{T-1}$ $(i = 1, \ldots, N)]$, and N technical indicators $[\{S^*_{i,t}\}_{t=0}^{T-1}$ $(i = 1, \ldots, N)]$, we estimate the slope coefficients and corresponding heteroskedasticity-robust t-statistics for the bivariate predictive regressions given by (1) and (10) in the paper for each i, as well as the principal component predictive regressions given by (11), (12), and (13) in the paper. Note that we compute the principal components in (11), (12), and (13) using $\{x^*_{i,t}\}_{t=0}^{T-1}$ and $\{S^*_{i,t}\}_{t=0}^{T-1}$ $(i = 1, \ldots, N)$, so that the pseudo sample accounts for the estimated regressors in the principal component predictive regressions. We store the t-statistics for all of the predictive regressions. Repeating this process 2,000 times yields empirical distributions for each of the t-statistics. For a given t-statistic, the empirical p-value is the proportion of the bootstrapped t-statistics greater than the t-statistic for the original sample.

Table A1 Predictive Regression Structural Stability Test Results, 1951:01–2011:12

Predictor	$\widehat{qLL}$	Predictor	$\widehat{qLL}$	Model	$\widehat{qLL}$
DP	−9.85	MA(1,9)	−10.02	PC-ECON	−21.54
DY	−10.23	MA(1,12)	−7.62	PC-TECH	−18.24
EP	−10.08	MA(2,9)	−7.02	PC-ALL	−6.88
DE	−7.08	MA(2,12)	−6.86		
RVOL	−11.92	MA(3,9)	−7.53		
BM	−9.88	MA(3,12)	−8.05		
NTIS	-17.88^{***}	MOM(9)	−8.01		
TBL	−12.04	MOM(12)	−7.50		
LTY	−9.80	VOL(1,9)	−6.26		
LTR	−8.77	VOL(1,12)	−7.30		
TMS	-13.24^{*}	VOL(2,9)	−8.57		
DFY	-13.89^{*}	VOL(2,12)	−9.29		
DFR	−8.76	VOL(3,9)	−8.63		
INFL	-14.67^{**}	VOL(3,12)	−8.75		

Notes. The second and fourth columns report the Elliott and Müller (2006) $\widehat{qLL}$ statistic for the bivariate predictive regression model,

$$r_{t+1} = \alpha_i + \beta_i q_{i,t} + \varepsilon_{i,t+1},$$

where r_{t+1} is the log equity risk premium (in percent) and $q_{i,t}$ is one of the 14 macroeconomic variables (14 technical indicators) given in the first (third) column. The last column reports the $\widehat{qLL}$ statistic for a predictive regression model based on principal components,

$$r_{t+1} = \alpha + \sum_{k=1}^{K} \beta_k \hat{F}_{k,t}^{j} + \varepsilon_{t+1},$$

where $\hat{F}_{k,t}^{j}$ is the kth principal component extracted from the 14 macroeconomic variables (j = ECON, PC-ECON), 14 technical indicators (j = TECH, PC-TECH), or the 14 macroeconomic variables and 14 technical indicators taken together (j = ALL, PC-ALL); $k = 3$ for PC-ECON, $k = 1$ for PC-TECH, and $k = 4$ for PC-ALL. The $\widehat{qLL}$ statistic is for testing the null hypothesis that the intercept and slope coefficients are constant. The 10%, 5%, and 1% critical values for the $\widehat{qLL}$ statistics in the second and fourth columns and for the PC-TECH model are −12.80, −14.32, and −17.57, respectively; the 10%, 5%, and 1% critical values for the $\widehat{qLL}$ statistic for the PC-ECON (PC-ALL) model are −23.37, −25.28, and −29.18 (−28.55, −30.60, and −35.09), respectively (where we reject for small values); *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A2 Predictive Regression Estimation Results for the Sentiment-Changes Index, 1965:08 to 2010:12

Macroeconomic variables					Technical indicators				
Predictor	Slope coefficient	R^2	R^2_{EXP}	R^2_{REC}	Predictor	Slope coefficient	R^2	R^2_{EXP}	R^2_{REC}
Panel A: Bivariate Predictive Regressions									
DP	0.12 [0.92]	0.26%	0.38%	−0.25%	MA(1,9)	0.20 [2.26]**	0.94%	−0.04%	4.94%
DY	0.15 [1.10]	0.37%	0.45%	0.08%	MA(1,12)	0.24 [2.48]***	1.19%	0.24%	5.07%
EP	0.09 [0.83]	0.17%	0.51%	−1.25%	MA(2,9)	0.16 [1.75]**	0.56%	−0.11%	3.27%
DE	0.02 [0.15]	0.00%	−0.05%	0.23%	MA(2,12)	0.26 [2.83]***	1.47%	0.62%	4.95%
RVOL	−0.41 [−0.55]	0.05%	0.16%	−0.40%	MA(3,9)	0.11 [1.25]	0.28%	−0.22%	2.31%
BM	0.19 [1.06]	0.27%	0.55%	−0.86%	MA(3,12)	0.14 [1.50]*	0.42%	0.06%	1.87%
NTIS	−0.27 [−0.14]	0.00%	0.05%	−0.18%	MOM(9)	0.19 [2.01]**	0.75%	0.35%	2.35%
TBL	0.02 [1.13]	0.20%	−0.13%	1.56%	MOM(12)	0.20 [2.15]**	0.84%	0.40%	2.64%
LTY	0.02 [1.20]	0.24%	0.25%	0.22%	VOL(1,9)	0.17 [1.88]**	0.63%	−0.02%	3.27%
LTR	−0.01 [−0.63]	0.07%	0.04%	0.17%	VOL(1,12)	0.18 [1.93]**	0.70%	0.14%	2.97%
TMS	0.01 [0.23]	0.01%	−0.13%	0.57%	VOL(2,9)	0.19 [2.10]*	0.86%	0.31%	3.07%
DFY	0.07 [0.77]	0.12%	−0.24%	1.59%	VOL(2,12)	0.20 [2.18]**	0.92%	0.57%	2.35%
DFR	−0.04 [−1.15]	0.30%	0.76%	−1.56%	VOL(3,9)	0.12 [1.23]	0.31%	0.10%	1.16%
INFL	0.05 [0.45]	0.04%	−0.08%	0.51%	VOL(3,12)	0.15 [1.54]*	0.47%	0.19%	1.60%
Panel B: Principal Component Predictive Regressions									
$\hat{F}_1^{ECON}$	0.01 [0.42]	0.07%	0.00%	0.38%	$\hat{F}_1^{TECH}$	0.03 [2.31]***	0.96%	0.28%	3.72%
$\hat{F}_2^{ECON}$	0.01 [0.32]								
$\hat{F}_3^{ECON}$	0.01 [0.25]								
Panel C: Principal Component Predictive Regression, All Predictors Taken Together									
$\hat{F}_1^{ALL}$	0.03 [2.15]**	1.11%	0.10%	5.21%					
$\hat{F}_2^{ALL}$	0.02 [0.93]								
$\hat{F}_3^{ALL}$	0.02 [0.61]								
$\hat{F}_4^{ALL}$	0.01 [0.26]								

Notes. Panel A reports estimation results for the bivariate predictive regression model,

$$\Delta\text{SENT}_{t+1} = \alpha_i + \beta_i q_{i,t} + \varepsilon_{i,t+1},$$

where ΔSENT_{t+1} is the sentiment-changes index and $q_{i,t}$ is one of the 14 macroeconomic variables (14 technical indicators) given in the first (sixth) column. Panels B and C report estimation results for a predictive regression model based on principal components,

$$\Delta\text{SENT}_{t+1} = \alpha + \textstyle\sum_{k=1}^{K} \beta_k \hat{F}_{k,t}^{j} + \varepsilon_{t+1},$$

where $\hat{F}_{k,t}^{j}$ is the kth principal component extracted from the 14 macroeconomic variables (j = ECON), 14 technical indicators (j = TECH), or the 14 macroeconomic variables and 14 technical indicators taken together (j = ALL). The brackets to the immediate right of the estimated slope coefficients report heteroskedasticity-consistent t-statistics; *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, based on one-sided (upper-tail) wild bootstrapped p-values; 0.00 indicates less than 0.005 in absolute value. The R^2 statistics in the third and eighth columns are computed for the full sample. The R^2_{EXP} (R^2_{REC}) statistics in the fourth and ninth (fifth and tenth) columns are computed for NBER-dated business-cycle expansions (recessions), as given by (9) in the paper.

Table A3 Tests of Exclusion Restrictions for the Conditional Fama-French Three-Factor Model of Momentum Portfolio Excess Returns, 1951:01–2011:12

Portfolio (i)	$\alpha_{i,1} = 0$, $\beta_{i,1}^{j} = 0\ \forall\ j$	$\alpha_{i,k} = 0$ for $k = 2, 3, 4$; $\beta_{i,2}^{j} = 0\ \forall\ j$; $\beta_{i,3}^{j} = 0\ \forall\ j$; $\beta_{i,4}^{j} = 0\ \forall\ j$	$\alpha_{i,k} = 0\ \forall\ k$, $\beta_{i,k}^{j} = 0\ \forall\ j, k$
# restrictions	4	12	16
Low	58.96***	66.61***	128.28***
2	72.67***	40.72***	117.87***
3	66.93***	39.21***	105.83***
4	50.71***	61.43***	118.21***
5	10.84**	23.77*	31.00*
6	4.15	50.31***	66.02***
7	26.70***	76.00***	111.32***
8	58.11***	113.34***	171.00***
9	58.30***	74.93***	112.46***
High	59.87***	41.63***	109.26***
UMD	108.15***	62.36***	156.86***

Notes. The table reports heteroskedasticity-consistent χ^2-statistics for tests of parameter restrictions for the conditional Fama-French three-factor model:

$$R_{i,t+1} - R_{f,t+1} = \alpha_{i,t} + \beta_{i,t}^{\text{MKT}}\text{MKT}_{t+1} + \beta_{,,t}^{\text{SMB}}\text{SMB}_{t+1} + \beta_{i,t}^{\text{HML}}\text{HML}_{t+1} + \varepsilon_{i,t+1},$$

where $R_{i,t+1}$ is the (simple) return on momentum portfolio i, $R_{f,t+1}$ is the risk-free return, MKT is the excess market return, SMB (HML) is the size (value) premium, and

$$\alpha_{i,t} = \alpha_{i,0} + \textstyle\sum_{k=1}^{4} \alpha_{i,k} \hat{F}_{k,t}^{\text{ALL}},$$

$$\beta_{i,t}^{j} = \beta_{i,0}^{j} + \textstyle\sum_{k=1}^{4} \beta_{i,k}^{j} \hat{F}_{k,t}^{\text{ALL}} \text{ for } j = \text{MKT}, \text{SMB}, \text{HML},$$

where $\hat{F}_{k,t}^{\text{ALL}}$ is the kth principal component extracted from 14 macroeconomic variables and 14 technical indicators taken together. The momentum portfolios are from Kenneth French's Data Library and formed from NYSE prior return deciles; UMD is the "up-minus-down" portfolio. The χ^2-statistics correspond to the exclusion restrictions given in the column heading; *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, based on wild-bootstrapped p-values.